Large subgraphs without complete bipartite graphs

David Conlon* Jacob Fox † Benny Sudakov ‡

Abstract

In this note, we answer the following question of Foucaud, Krivelevich and Perarnau. What is the size of the largest $K_{r,s}$-free subgraph one can guarantee in every graph G with m edges? We also discuss the analogous problem for hypergraphs.

1 Introduction

Motivated by the classical Turán problem, Foucaud, Krivelevich and Perarnau [3] proposed to study the size of the largest H-free subgraph one can always find in every graph G with m edges. Denote this function by $f(m, H)$. It is easy to determine $f(m, H)$ asymptotically if H is not bipartite. In [3], the authors studied this problem when forbidding all even cycles in the subgraph up to length $2k$ and obtained estimates that are tight up to a logarithmic factor. They also asked to determine $f(m, H)$ when H is a complete bipartite graph. The goal of this note is to resolve this question.

2 Complete bipartite graphs

Let $K_{r,s}$ be the complete bipartite graph with parts of order r and s, where $2 \leq r \leq s$. The following theorem gives a lower bound on $f(m, K_{r,s})$.

Theorem 2.1. *Every graph G with m edges contains a $K_{r,r}$-free subgraph of size at least $\frac{1}{4}m^{\frac{r}{r+1}}$.*

To prove this theorem we need un upper bound on the maximum number of copies of $K_{r,r}$ which one can find in a graph with m edges. The problem of maximizing the number of copies of a fixed graph H was solved by Alon [1] for all graphs and by Friedgut and Kahn [4] for all hypergraphs. For our purposes the following easy estimate will suffice.

Lemma 2.2. *Every graph G with m edges contains at most $2m^r$ copies of $K_{r,r}$.*

Proof. Note that every copy of $K_{r,r}$ in G contains a matching of size r. Clearly the number of such matchings in G is at most $\binom{m}{r}$. Also note that every matching in G of size r can appear in at most 2^r copies of $K_{r,r}$. This implies that the total number of such copies is at most $2^r\binom{m}{r} \leq 2m^r$. □

Using this lemma, together with a simple probabilistic argument, one can prove a lower bound on $f(m, K_{r,s})$.

*Mathematical Institute, Oxford OX2 6GG, UK. Email: david.conlon@maths.ox.ac.uk.
†Department of Mathematics, MIT, Cambridge, MA 02139-4307. Email: fox@math.mit.edu.
‡Department of Mathematics, ETH, 8092 Zurich, Switzerland and Department of Mathematics, UCLA, Los Angeles, CA 90095. Email: benjamin.sudakov@math.ethz.ch.

Proof of Theorem 2.1. Let G be a graph with m edges. Consider a random subgraph G' of G, obtained by choosing every edge randomly and independently with probability $p = \frac{1}{2}m^{-1/(r+1)}$. Then the expected number of edges in G' is mp. Also, by Lemma 2.2, the expected number of copies of $K_{r,r}$ in G' is at most $2p^{r^2}m^r$. Delete one edge from every copy of $K_{r,r}$ contained in G'. This gives a $K_{r,r}$-free subgraph of G, which by linearity of expectation, has at least

$$pm - 2p^{r^2}m^r \geq \frac{1}{2}m^{\frac{r}{r+1}} - \frac{1}{8}m^{\frac{r}{r+1}} \geq \frac{1}{4}m^{\frac{r}{r+1}}$$

edges on average. Hence, there exists a choice of G' which produces a $K_{r,r}$-free subgraph of G of size at least $\frac{1}{4}m^{\frac{r}{r+1}}$. □

Next we show that this gives an estimate on $f(m, K_{r,s})$ which is tight up to a constant factor depending on s by taking G to be an appropriately chosen complete bipartite graph with m edges.

Theorem 2.3. *Let $2 \leq r \leq s$ and let G be a complete bipartite graph with parts U and W, where $|U| = m^{1/(r+1)}$ and $|W| = m^{r/(r+1)}$. Then G has m edges and the largest $K_{r,s}$-free subgraph of G has at most $sm^{r/(r+1)}$ edges.*

Proof. The proof is a simple application of the counting argument of Kővári-Sós-Turán [5]. Let G' be a $K_{r,s}$-free subgraph of G and let $d = e(G')/|W|$ be the average degree of vertices of G' in W. If $d \geq s$, then, by convexity,

$$\sum_{w \in W} \binom{d_{G'}(w)}{r} \geq |W|\binom{d}{r} \geq \binom{s}{r} m^{r/(r+1)} \geq sm^{r/(r+1)}/r!\,.$$

On the other hand, since G' is $K_{r,s}$-free we have that

$$\sum_{w \in W} \binom{d_{G'}(w)}{r} < s\binom{|U|}{r} \leq s|U|^r/r! = sm^{r/(r+1)}/r!\,.$$

This contradiction completes the proof of the theorem. □

Remarks.

- Since $K_{2,2}$ is also a 4-cycle, our result improves by a logarithmic factor an estimate obtained by Foucaud, Krivelevich and Perarnau [3].
- Since the Turán number for $K_{r,s}$ is not known in general, it is somewhat surprising that one can prove a tight bound on the size of the largest $K_{r,s}$-free subgraph in graphs with m edges.

3 Hypergraphs

The results presented in the previous section can be extended to k-uniform hypergraphs, which, for brevity, we call k-graphs. Given a fixed k-graph H, let $f(m, H)$ denote the size of the largest H-free subgraph one can always find in every k-graph G with m edges. Let $K^{(k)}_{r,\dots,r}$ denote the complete k-partite k-graph with parts of size r.

Theorem 3.1. *Every k-graph G with m edges contains a $K^{(k)}_{r,\dots,r}$-free subgraph of size at least $\frac{1}{4}m^{\frac{q-1}{q}}$, where $q = \frac{r^k-1}{r-1}$.*

Proof. Let G be a k-graph with m edges. Every copy of $K^{(k)}_{r,\ldots,r}$ in G contains a matching of size r and the number of such matchings is at most $\binom{m}{r}$. On the other hand, every matching in G of size r can appear in at most $(k!)^r$ copies of $K_{r,r}$. This implies that the total number of such copies is at most $(k!)^r\binom{m}{r}$.

Consider a random subgraph G' of G, obtained by choosing every edge randomly and independently with probability $p = \frac{1}{2}m^{-1/q}$. Then the expected number of edges in G' is mp and the expected number of copies of $K^{(k)}_{r,\ldots,r}$ in G' is at most $(k!)^r p^{r^k}\binom{m}{r}$. Delete one edge from every copy of $K^{(k)}_{r,\ldots,r}$ contained in G'. This gives a $K^{(k)}_{r,\ldots,r}$-free subgraph of G with at least

$$pm - (k!)^r p^{r^k}\binom{m}{r} \geq \frac{1}{4}m^{\frac{q-1}{q}}$$

expected edges. Hence, there exists a choice of G' which produces a $K^{(k)}_{r,\ldots,r}$-free subgraph of G of this size. □

We can again see that this estimate is tight up to a constant factor depending on r.

Theorem 3.2. *Let $2 \leq r, k$, $q = \frac{r^k-1}{r-1}$ and let G be a complete k-partite k-graph with parts $U_i, 1 \leq i \leq k$, such that $|U_i| = m^{r^{i-1}/q}$. Then G has m edges and the largest $K^{(k)}_{r,\ldots,r}$-free subgraph of G has at most $rm^{(q-1)/q}$ edges.*

The proof of this theorem uses a similar counting argument to the graph case but is more involved. It follows from the following statement, which one can prove by induction. This technique has its origins in a paper of Erdős [2].

Proposition 3.3. *Let G be a k-partite k-graph with parts $U_i, 1 \leq i \leq k$, such that $|U_i| = n^{r^i}$ and with $a\prod_{i\geq 2}|U_i|$ edges and $a \geq r$. Then G contains at least $\binom{a}{r}\prod_{i\leq k-1}\binom{|U_i|}{r}$ copies of $K^{(k)}_{r,\ldots,r}$.*

Proof. We prove this by induction on k. The base case $k = 1$ is trivial, by properly interpreting empty products as one.

Now suppose we know the statement for $k-1$. For every vertex $x \in U_k$, denote by G_x the $(k-1)$-partite $(k-1)$-graph which is the link of vertex x (i.e., the collection of all subsets of size $k-1$ which together with x form an edge of G). Let $a_x\prod_{i=2}^{k-1}|U_i|$ be the number of edges in G_x. By definition, $\sum_x a_x = a|U_k| = an^{r^k}$. By the induction hypothesis, each G_x contains at least $\binom{a_x}{r}\prod_{i\leq k-2}\binom{|U_i|}{r}$ copies of $K^{(k-1)}_{r,\ldots,r}$. By convexity, the total number of such copies added over all G_x is at least

$$\binom{a}{r}n^{r^k}\prod_{i\leq k-2}\binom{|U_i|}{r} = \binom{a}{r}|U_{k-1}|^r\prod_{i\leq k-2}\binom{|U_i|}{r} \geq r!\binom{a}{r}\prod_{i\leq k-1}\binom{|U_i|}{r} \geq a\prod_{i\leq k-1}\binom{|U_i|}{r}.$$

For every subset S which intersects each U_i with $i \leq k-1$ in exactly r vertices, denote by $d(S)$ the number of vertices $x \in U_k$ such that x forms an edge of G together with every subset of S of size $k-1$ which contain one vertex from every U_i. By the above discussion, we have that $\sum_S d(S) \geq a\prod_{i\leq k-1}\binom{|U_i|}{r}$, that is, at least the number of all copies of $K^{(k-1)}_{r,\ldots,r}$ in all G_x. On the other hand, by the definition of $d(S)$, the number of copies of $K^{(k)}_{r,\ldots,r}$ in G equals $\sum_S\binom{d(S)}{r}$. Since the total number of sets S is $\prod_{i\leq k-1}\binom{|U_i|}{r}$, the result now follows by convexity. □

Acknowledgment. We would like to thank M. Krivelevich for bringing this problem to our attention and for sharing with us preprint [3].

References

[1] N. Alon, On the number of subgraphs of prescribed type of graphs with a given number of edges, *Israel J. Math.* **38** (1981), 116–130.

[2] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.* **2** (1964), 183–190.

[3] F. Foucaud, M. Krivelevich and G. Perarnau, Large subgraphs without short cycles, preprint.

[4] E. Freidgut and J. Kahn, On the number of copies of one hypergraph in another, *Israel J. Math.* **105** (1998), 251–256.

[5] T. Kővári, V.T. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloq. Math.* **3** (1954), 50–57.